

The Impact of Expertise on Brain Computer Interface Based Salient Image Retrieval

Ashkan Yazdani¹, Jean-Marc Vesin², Dario Izzo³, Christos Ampatzis³, Touradj Ebrahimi¹

Abstract—Autonomous decision making modules in computer vision application allow recognition and classification of different objects, persons, and events in images and video sequences and also make it possible to classify different sensor readings (e.g. images) according to their scientific saliencies. In this paper, we propose a new approach to create the training set for these algorithms by retrieving salient images using electroencephalogram (EEG) and brain computer interface (BCI) and rapid image presentation. To this end, two groups of subjects, namely, expert and novice subjects were asked to participate in our experiments. We show that a relatively high retrieval accuracy can be achieved for most of the subjects. Furthermore, to assess the impact of expertise on the retrieval process, we study their EEG signals separately and show that there is a clear difference in their brainwaves while observing salient images.

I. INTRODUCTION

Designing new algorithms to classify different sensor readings (e.g. images) according to their level of scientific saliency is one of the ultimate goals of many automated computer vision and multimedia content analysis applications. An autonomous system capable of recognizing objects, events, and abnormality in images or video sequences can be used in various applications ranging from medicine to space and defence. The main difficulty, however, is to define what is scientifically interesting. Given a well-defined training set, machine learning algorithms could be trained directly to classify what is scientifically interesting and what is not, without further information about these two very broad classes.

A typical approach to create such training data is to interview or interrogate expert scientists on a particular set of images, and to ask them to simply classify or rank them so that a computer can be trained to have a response similar to that of the interviewed expert scientists. In this way, the computer has to automatically extract the relevant features that guided the expert's decision-making and learn to use them in such a way as to mimic the expert's classification. Despite the simplicity of this methodology, various drawbacks may be listed. For example, it requires the scientists to undergo long and time-consuming sessions of image classification that may prove to be particularly tiring and cumbersome, which in turn can result in the acquisition of a noisy training set. Moreover, this approach is subject to the fuzziness of

the scientist's reasoning when placing a highly cognitive judgment upon each picture. In other words, the scientist will repeatedly consciously filter the image, eventually merging even contradictory verdicts to one binary classification or ranking. In this work, we propose an alternative approach to creating such a training set for a classifier; in particular, the information about the expert's classification is extracted directly from the classification of his/her brainwaves. There have been several attempts to detect objects of interest contained in the visual stimuli presented to the subjects based on brain activity and content analysis.[1], [2], [3], [4]. It is well known from neurophysiological studies that when subjects look at images which arouse mental response such as surprise, anticipation, etc., their parietal cortex is excited in a very characteristic way: a synchronized peak in the global electrical activity of large groups of neurons in the parietal area arises after approximately 300 ms after the stimulus (image) presentation. This electrical activity can be recorded with an EEG instrument as an electric positive potential wave and is commonly referred to as P300 [5].

In this paper we explore the impact of expertise on the retrieval efficiency of salient images using a P300-based BCI system. We try to answer the question whether there is a difference in the brain activity of expert and novice subjects while watching a visual stimulus which is scientifically interesting for them.

The rest of this paper is organized as follows. In Sections II and III, the experiments performed and the signal processing and classification algorithms used in this study are described. A discussion of results follows in Section IV and finally Section V concludes the paper.

II. EXPERIMENTAL SETUP

The performed experiments aimed at investigating the relation between event related potentials (ERPs) and expert knowledge or scientific curiosity. In order to assess the impact of expertise on ERPs, an oddball paradigm was used and a special set of visual stimuli was presented rapidly to the subjects. Two types of experimental subjects, namely, a person who has profound scientific knowledge about the stimuli and novice subjects were asked to participate in the experiments.

The visual stimuli used in the experiments were taken from the European Space Agency's database of "multilayer coatings for thermal applications"¹. The database contains images obtained during the process of designing a

¹ Ecole Polytechnique Fédérale de Lausanne (EPFL), Institute of Electrical Engineering (IEL), Multimedia Signal Processing Group (MMSPG).

² Ecole Polytechnique Fédérale de Lausanne (EPFL), Institute of Electrical Engineering (IEL), Applied Signal Processing Group (ASPG).

³ European Space and Technology Research Center (ESTEC), Advanced Concepts Team (ACT).

¹The database can be visited at the link www.esa.int/gsp/ACT/nan/op/bigrunresults.htm

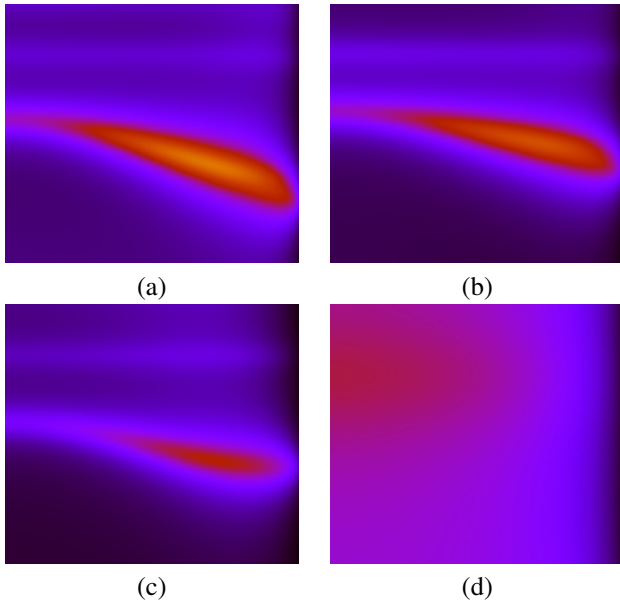


Fig. 1. Examples of a target (a), obvious oddball (b), non-obvious oddball (c) and standard (d) images used in the experiment.

multilayered material exhibiting predefined thermal emissivity profiles (which are called *targets*). Spectral directional properties of a material can be presented as 2-dimensional contour plots with axes representing angle and wavelength parameters and colour-coded with respect to the magnitude of the target parameter (for example emittance). Different materials, including the ideal target solution, correspond to different plots which appear as different 2-dimensional contours. However, since a material which exactly matches the desired properties is not obtainable, the best solution found will be only to a certain degree similar to the ideal target solution. This “degree of similarity” is related to a simple pattern matching process (e.g. the image looks similar to the target image) for novice subjects, and to more complex cognitive processes in the expert.

This experiments aimed at answering to the question whether there is a difference in P300 responses between subjects who possess scientific knowledge about presented stimuli and novice subjects. The experiment used a modification of the oddball paradigm, with two types of oddballs: obvious and non-obvious. At the beginning of each session, an image corresponding to the target solution (see Figure 1 (a)) was presented to the subjects and they were instructed to “look for similar images” in the stimuli presented and silently count how many targets they saw. Then a sequence of images was presented, which contained plots of materials with properties different from the ideal target (standard images), very similar to the target (obvious oddballs) and slightly similar to the target (non-obvious oddballs). Examples of such images are shown in Figure 1, whilst the parameters of the experiment are presented in Table I.

In total, five subjects were used, including one expert (the European Space Agency’s scientist conducting the aforementioned study on multilayered materials (represented as

TABLE I
PARAMETERS OF THE PERFORMED EXPERIMENT

No. of subjects	No. of targets	No. of sequences per target	Images in seq.	Oddballs in seq.	Repetitions	IDP/IIP (ms)	T (s)
4+1	2	5	50	3+3	2	500/0	25

subject 1) and four novices (subjects 2, 3, 4, and 5). Two different target images were used, with five image sequences prepared for each of them. Every sequence contained three obvious and three non-obvious oddballs. Each measurement was conducted twice in two different acquisition sessions which were no longer than 30 minutes from each other. A moderately fast image presentation rate with image display period (IDP) of 500ms and without inter-image period (IIP) was used, which resulted in sequences of duration 25 seconds in length.

III. SIGNAL PROCESSING AND MACHINE LEARNING

In each experiment, the image sequences were presented to the subjects and the exact system clock of each image presentation onset was stored. The EEG signal was recorded synchronously with image presentation at a 2048 Hz sampling rate from 32 active electrodes placed at the standard positions of the 10-20 international system. A Biosemi Active Two amplifier was used for amplification and 24-bit analog to digital conversion. After pre-amplification and digitalization of the EEG signal, a block of preprocessing filters was applied to the raw signals to remove the various artifacts. Wavelet decomposition was then applied to break down the signal into different frequency components and for further multiresolution analysis. In the next step, features were extracted from the signal and finally the extracted feature vectors were classified.

For preprocessing, the following methods were applied to the raw signal. First, the average signal from all electrodes was used for referencing and a 12th order forward-backward Butterworth bandpass filter with zero phase shift and the cut-off frequencies of 1 Hz and 95 Hz was used to remove all high frequency non-EEG artifacts as well as low frequency drifts from the recorded signal. In the next step, the EEG was downsampled from 2048 Hz to 256 Hz using an eight-order Lowpass Chebyshev Type I filter in order to reduce the number of samples of the signal.

After this preprocessing phase, discrete wavelet transform (DWT) was used to decompose the EEG signal into its subbands. Wavelet transform is known to be effective for representing various aspects of signals such as trends, discontinuities, and repeated patterns when other signal processing approaches fail or are not as effective. The EEG signal contains non-stationary components and it has been shown that wavelets clearly outperform Fourier transform based methods for feature extraction from the EEG signals [6]. The Daubechies family of wavelets is one of the most commonly used orthogonal ones satisfying the admissibility conditions, thus allowing reconstruction of the original signal from the wavelet coefficients. Daubechies wavelets of different

orders were investigated for the analysis of epileptic EEG records. Daubechies of order 4 and higher were found to be appropriate tools for analysis of EEG data [7]. In this study, Daubechies order 6 was used for the decomposition of EEG signals into different bands. Figure 2 illustrates a typical EEG signal used in this study and its decomposition using DWT. It has been shown in [6] that only the delta and theta frequencies play a major role in the P300 response. However it was discussed that alpha frequencies have also minor roles. Therefore, these three subbands (the last four rows in Figure 2, which correspond to 1-2Hz, 2-4Hz, 4-8 Hz, and 8-16 Hz) were summed together to reconstruct the P300 signal.

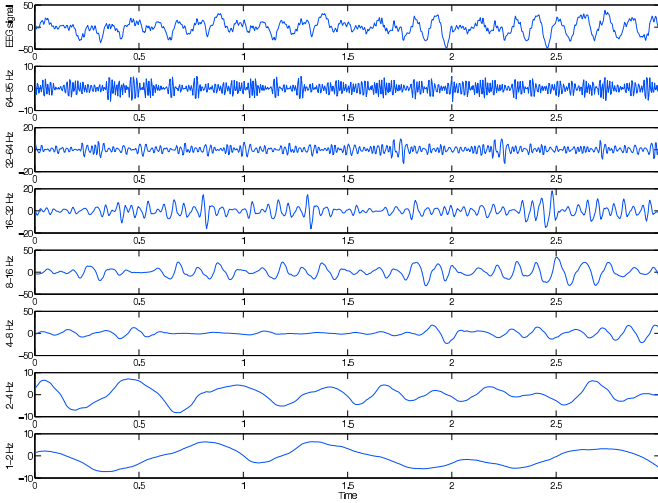


Fig. 2. Three second long EEG signal and its decomposition

In the next step, single trials of duration 1000 ms were extracted from the data. Single trials started at stimulus onset (provided by the image presentation software), i.e. at the beginning of the intensification of the first image in the image sequence, and ended 1000 ms after stimulus onset. The EEG signal was again downsampled from 256 Hz to 32 Hz using an eight-order Lowpass Chebyshev Type I filter. Therefore, the decimated signal contains frequency components of 1-12.8 Hz and each single trial contains 32 samples. 16 electrodes (central, parietal and occipital) were selected and the signals of these electrodes were concatenated to form a feature vector corresponding to that single trial.

Eyeblinks, eye movements, muscle activity, or subject movement can cause large amplitude outliers in the EEG. To reduce the effects of such outliers, the data from each electrode were windosized. For the samples from each electrode the 10th percentile and the 90th percentile were computed. Amplitude values lying below the 10th percentile or above the 90th percentile were then replaced by the 10th percentile or the 90th percentile, respectively. A support vector machine (SVM) classifier with radial basis function kernels was used for the classification of single trials. To perform the classification using this method, five-fold cross validation was performed and it was repeated 15 times. In each training run, the extracted feature vectors were first

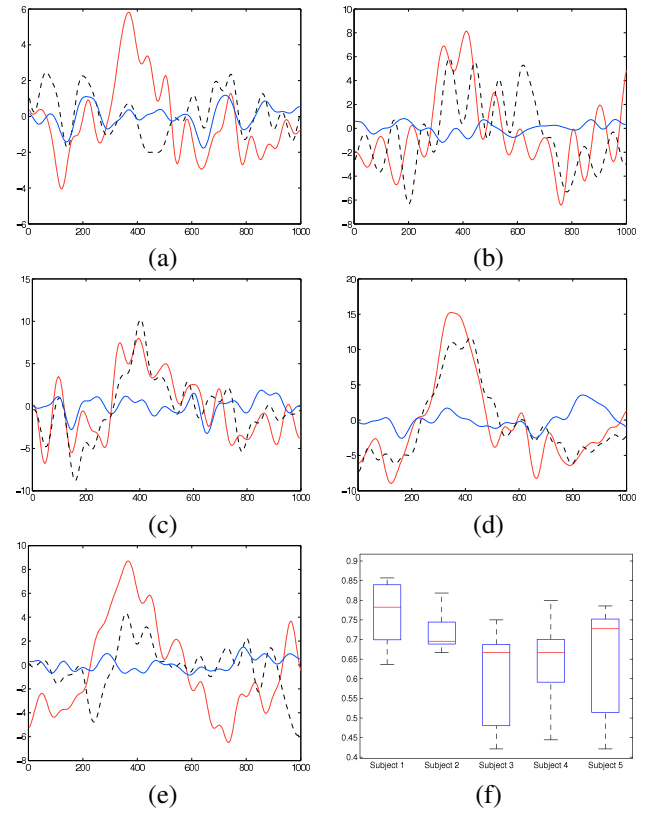


Fig. 3. The averaged obvious target (red), non-obvious target (dashed black) and nontarget (blue) EEG signals for different subjects (subject 1 (a) to subject 5 (e)). The boxplot of F-measure values obtained in single trial classification of obvious and non-obvious oddball signals for different subjects (f).

preprocessed using principal component analysis (PCA) and the dimension of the feature vectors was reduced from 512 to 288. The projection matrix was used to be applied later for dimension reduction on test data. In the next step, the training and test data were normalized to have the minimum and maximum values of zero and one, respectively. Finally, the training data was used to train the SVM classifier.

IV. RESULTS

In this section, the results on averaged signal analysis and classification of single trials are presented and discussed.

In order to analyze the P300 signal, the obvious oddball, non-obvious oddball, and non-target (standard) signals were averaged separately for each subject. Figure 3 illustrates these averaged signals, taken from the PZ electrode for different subjects. As it can be observed, most of the subjects are unable to distinguish among the obvious and non-obvious oddballs. This can be inferred from Figure 3 in that for most of the subjects there is a clear P300 peak occurring in their EEG signals even when they observe non-obvious oddballs. However, the amplitude of the two P300 signals (corresponding to obvious and non-obvious oddballs) might be different, as the variety of non-obvious targets is large and hence it is also possible that the subjects classify some of the non-obvious oddballs as non-targets. Interestingly, the expert was the only subject who could distinguish fully between the

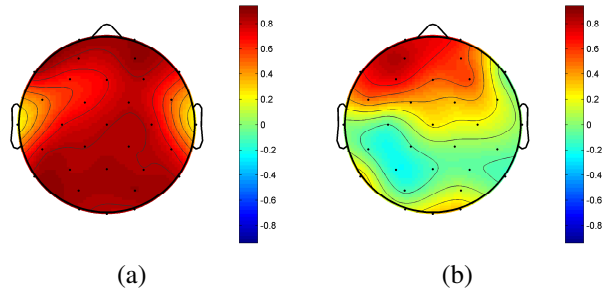


Fig. 4. The correlation of averaged obvious and non-obvious oddball signals of different electrodes for novice subjects (a) and expert (b).

two sub-category of target images and hence there existed no trace of P300 in his brainwaves when observing non-obvious oddballs and he simply classified them as non-targets. Figure 4 illustrates the correlation coefficient values of the averaged obvious and non-obvious oddball signals in different parts of the scalp. As it can be seen, for novice subjects a high correlation value exists between the two signals in almost all electrodes, whereas for the expert subject there is no significant correlation in the central, parietal, and occipital electrodes, where the P300 signal is mostly expected to emerge.

The results of single trial classification of obvious and non-obvious oddball signals is presented in Figure 3 (f). Once again it can be observed that for the expert subject, the highest F-measure among all subjects is obtained. Furthermore, it can be seen that for most of the novice subjects, the results of the classification can be sometimes very close or even inferior to that of the random guess for some test datasets during cross validation. Table II presents the single trial classification results of oddball (both obvious and non-obvious) and non-target signals. It can be seen that, using the aforementioned BCI system, relatively high values of F-measure can be obtained for the salient (scientifically interesting) images retrieval for most of the subjects. However, the accuracy of the retrieval system based on P300 and BCI can vary among subjects, depending on various physiological and affective factors. The classification results for the expert subject is lower in this case compared to novice subjects. One reason could be that the non-obvious oddball signals which are labeled as oddballs are quite similar with non-targets and hence the trained classifier can not separate them correctly which results in a low recall rate. As mentioned earlier, the amplitude of P300 can vary among subjects due to their age, fatigue, level of interest, and other factors [8], [9], [10]. As it can be seen in Figure 3(a), the P300 amplitude for the only expert subject used in this study is lowest among all subjects. This also can influence the relatively lower recall rate and F-measure obtained for the expert subject compared to others.

V. CONCLUSIONS AND FUTURE WORKS

In this paper, the feasibility of P300-based BCI system for retrieving salient images was studied. We showed that for all subjects relatively high retrieval accuracies can be obtained.

TABLE II
ODDBALL VS. NON-TARGET CLASSIFICATION RESULTS

subject	F-measure	precision	recall
1	56.4 ± 15.6	83.0 ± 12.6	43.3 ± 17.3
2	71.1 ± 9.0	79.1 ± 10.5	67.33 ± 15.0
3	56.8 ± 15.6	76.7 ± 14.8	46.7 ± 21.9
4	69.6 ± 7.7	79.0 ± 13.0	65.3 ± 14.4
5	35.0 ± 10.4	78.7 ± 18.7	18.3 ± 11.3

Furthermore, the differences in the brainwaves of the expert and novices during the retrieval process was studied. We showed that the expert was the only subject who could fully distinguish obvious and non-obvious oddballs. Consequently his brain activity while watching non-obvious oddballs was quite similar to that of when watching non-target images.

Several aspects of this work can be further investigated. First, a larger number of subjects (both experts and novices) should be enrolled in the experiments to investigate the possibility to generalize the results obtained in this study. Moreover, in this study, the subjects were told to look for a certain target pattern in the stimuli. However, this can be relaxed and the subjects can be asked to find whichever pattern they find interesting. This way, it can be possible to assess the detection of scientific curiosity.

VI. ACKNOWLEDGMENTS

The authors wish to acknowledge Swiss National Science Foundation grant no. 116253 and the European Space Agency Ariadna scheme for having initiated and supported this research.

REFERENCES

- [1] N. Bigdely-Shamlo, A. Vankov, R. Ramirez, and S. Makeig, "Brain activity-based image classification from rapid serial visual presentation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 5, pp. 432–441, 2008.
- [2] A. Kapoor, P. Shenoy, and D. Tan, "Combining brain computer interfaces with vision for object categorization," in *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008, 2008*, pp. 1–8.
- [3] L. Parra, C. Christoforou, A. Gerson, M. Dyrholm, A. Luo, M. Wagner, M. Philastides, and P. Sajda, "Spatiotemporal linear decoding of brain state: Application to performance augmentation in high-throughput tasks," *IEEE, Signal Processing Magazine*, vol. 25, no. 1, pp. 95–115, 2008.
- [4] J. Wang, E. Pohlmeier, B. Hanna, Y. Jiang, P. Sajda, and S. Chang, "Brain state decoding for rapid image retrieval," in *Proceedings of the seventeen ACM international conference on Multimedia*. ACM, 2009, pp. 945–954.
- [5] T. Picton, "The P300 wave of the human event-related potential," *Journal of clinical neurophysiology*, vol. 9, no. 4, p. 456, 1992.
- [6] V. Kolev, T. Demiralp, J. Yordanova, A. Ademoglu, and Ü. Isoglu-Alkaç, "Time-frequency analysis reveals multiple functional components during oddball P300," *NeuroReport*, vol. 8, no. 8, p. 2061, 1997.
- [7] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084–1093, 2007.
- [8] C. Kraiuhin, E. Gordon, P. Stanfield, R. Meares, and A. Howson, "P300 and the effects of aging: Relevance to the diagnosis of dementia," *Experimental Aging Research*, vol. 12, no. 4, pp. 187–192, 1986.
- [9] L. Hoffman and J. Polich, "P300, handedness, and corpus callosal size: gender, modality, and task," *International Journal of Psychophysiology*, vol. 31, no. 2, pp. 163–174, 1999.
- [10] J. Polich, "P300 clinical utility and control of variability," *Journal of Clinical Neurophysiology*, vol. 15, no. 1, p. 14, 1998.